

Bandolier 102

Independent evidence-based health care

A new beginning

In cricket we are told that batsmen who reach a century should calm down, take a fresh guard, and re-set their minds so that they start on another hundred. So a new beginning for **Bandolier** might have been better with issue 101, but 102 is pretty close.

The good news is that a lot of individuals and organisations have signed up to the new regime, and the **Bandolier** print run will certainly be healthy, if not quite as large as before. So thank you to all who have subscribed, and helped keep **Bandolier** running as an independent source of evidence-based information on healthcare. What we also want to continue is the regular flow of questions and queries from readers. We may not be able to answer them all immediately, but they inform our thinking and searching, and keep us relevant.

This month

Studies keep coming out about how much evidence is being used, and sometimes about how much it is not used. This month we were taken with a study indicating that guidelines can be evidence-free zones, or at least limited in the amount of evidence they quote in the form of randomised trials and systematic reviews. Whenever people look at guidelines, problems about the use of evidence arise.

Bandolier is also interested in things we do that don't work. This month a study of arthroscopy for osteoarthritic knees shows that, whatever else, lavage and arthroscopy relieve pain no more than a placebo operation. In the USA, that is a \$3.5 billion dollar industry. We are also interested in evidence that for-profit hospitals have higher death rates than not-for-profit, the links between smoking and impotence, family interventions for delinquency, depression in advanced disease, and a re-assertion on the efficacy of influenza vaccination in older people.

Internet Bandolier

There is much new, especially on pain. The acute pain ladder has been updated following updates on a number of common analgesics, as has the ladder for migraine treatments. There is also more evidence on paracetamol and coxibs in arthritis. Most of the useful papers on gout that we could find have been put onto the site. And finally, there is a review of reviews on sildenafil for erectile dysfunction.

EVIDENCE-BASED GUIDELINES?

Bandolier is interested in guidelines and their development. On the face of it they seem like a good idea because they should condense all the best knowledge and experience to give any individual practitioner the confidence that, within limits, they can aspire to the same level of practice as the best in their field. **Bandolier's** Internet pages have examples of good guidelines [1].

Variability of guidelines

Not all guidelines are created equal and there are examples of great variability between the advice of guidelines. One revelation from Newcastle was that of guidelines for anti-coagulation for atrial fibrillation in the UK [2].

In 1996 various people and organisations in England, Wales and Scotland were contacted about the existence of guidelines for anticoagulant treatment of atrial fibrillation. These included regional and national NHS bodies, professional and charitable institutions, and members of mailing lists of audit organisations. They represented purchasers and providers of healthcare and relevant national organisations.

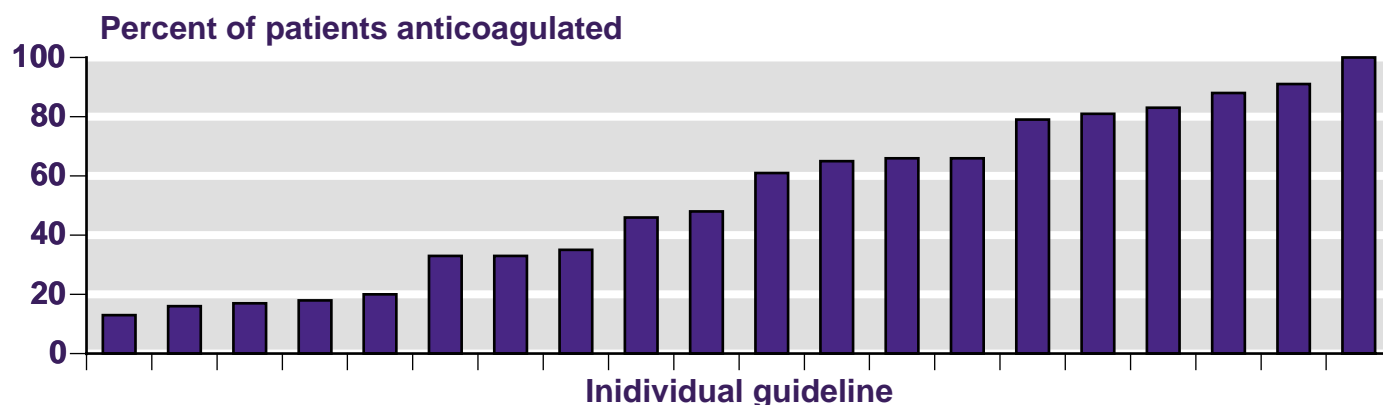
Guidelines were defined as a document produced to help clinicians decide which patients should be given anticoagulant drugs. Drafts, or documents designed for single specialised units, or to provide guidance once warfarin treatment had begun were not included. Where possible, guideline developers were interviewed using a semistructured method about how guidelines had been developed.

All included guidelines were applied to 100 consecutive patients with atrial fibrillation aged 65 years or older identified in a community survey. Details of risk factors for stroke or contraindications for treatment were obtained.

In this issue

Evidence-based guidelines.....	p. 1
Surgery for arthritic knees	p. 3
Family/parenting interventions for delinquency	p. 4
Smoking and impotence	p. 5
Death and Profit	p. 6
Depression in advanced disease	p. 7
Influenza vaccine in over-65s	p. 8

Figure 1: Anticoagulation guidelines in the UK applied to the same 100 consecutive patients



Results

The overall response rate was 66% (350/534), yielding 48 documents of which 20 fulfilled the requirements for definition of a guideline. They varied from a single page to 28 pages, were primarily for use by general practitioners, and affected populations from 12,000 to 500,000.

Guidelines were not systematically developed. A group of people developed about half, and a single person developed the other half. About half had some outside consultation, but about a quarter had no external review. Distribution was haphazard and few had educational meetings to introduce the guideline. Only one was explicitly claimed to be evidence-based, and had outside consultations from a health economist and clinician, with external review and local consultation, with wide distribution and an educational meeting to introduce the guideline.

When applied to 100 consecutive patients, the number recommended for anticoagulation by the guidelines ranged from 13 to 100 (Figure 1). Only one patient would have had anticoagulant treatment recommended by all guidelines, but every patient would have been recommended for anticoagulation by at least two guidelines (but not the same two). Target INR values varied between 1.2 to 1.5 and 2.5 to 3.0.

Evidence-base of guidelines

Another examination of the evidence-base of guidelines comes from Greece [3]. Researchers looked for guidelines published in 1979, 1984, 1989, 1994 and 1999 in six prestigious English language journals (Annals of Internal Medi-

cine, BMJ, JAMA, NEJM, Lancet and Pediatrics). The definition of what constituted a guideline was specific and included all articles containing the words guideline or recommendation or other characteristic words in title or abstract and which had a main focus on prevention or therapeutic interventions.

For each guideline, the reference lists were scrutinised and characterised as randomised trial, systematic review, meta-analysis, or neither. All cited articles were searched in MEDLINE and full records and abstracts were scrutinised. The full paper was retrieved and read in the case of any uncertainty about whether it was a randomised trial. This also applied to articles published before 1966.

Where guidelines had references, contained fewer than two citations of randomised trials, and cited no systematic reviews, a full MEDLINE search was performed up to the year of publication of the guideline.

Results

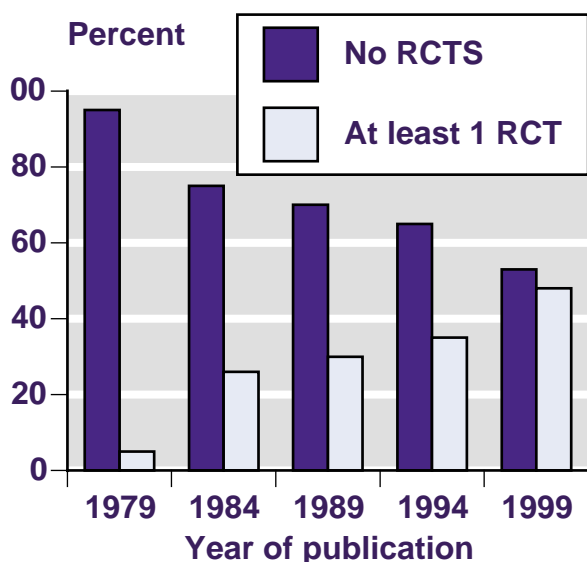
There were 191 guidelines identified in these six journals for the years searched, predominantly from the USA (86%). Group authorship was most common (84%). Of the 191 guidelines only 12 (6%) had performed a systematic review, but 130 (68%) made no mention about a lack of evidence. Thirty-six guidelines (19%) had no references.

Randomised trials made up a minority of the citations (Table 1). Only 8% of the citations were randomised trials, and fewer than 1% were systematic reviews or meta-analyses of randomised trials or epidemiological studies. The pro-

Table 1: Citations in 191 guidelines in six prestigious medical journals

Citations	Number	Percent of total citations	Mean citations per guideline
Total from 191 guidelines	4853	100	25.4
Randomised trial	393	8.1	2.1
Systematic review	19	0.4	0.1
Meta-analysis of randomised trials	23	0.5	0.1
Meta-analysis of epidemiological studies	11	0.2	0.1
Books/pamphlet/brochure	719	14.8	3.8
Abstract	122	2.5	0.6

Figure 2: Guidelines and their citation of RCTs



portion of guidelines not citing any randomised trials fell from 95% in 1979 to 53% in 1999 (Figure 2). Only about one guideline in 10 had a systematic review or meta-analysis.

Thirty-nine guidelines had fewer than two randomised citations with no systematic review. Because 30 were in Pediatrics, 10 of these were chosen at random, and 19 in all were the subject of specific searches. In 12 of the 19, additional relevant randomised trials were found. The number of additional randomised trials was 1 to 194 per topic.

Comment

Guidelines are important, and are proliferating. They can be individual, local, regional or national, and often many variants of the same guidelines exist at the same time. Guidelines should be updated regularly. Most importantly, they should be based on the best available evidence. The evidence is that many, perhaps most, are not.

These two papers look at a particular clinical condition exclusively in the UK, and more widely at material that turned out to be predominantly from the USA. They give the same answer, that when examined guidelines are not good enough. The Greek review [2] interestingly looks at parameters associated with citing randomised evidence in guidelines. Those funded by government and professional bodies were worse than those with university and private (usually pharmaceutical) sources of funds. The lesson is that we should not take any guideline on trust, without a sceptical examination of how it has been arrived at, and whether it has followed good practice for guideline development.

References:

- 1 R Thomson et al. Decision analysis and guidelines for anticoagulation therapy to prevent stroke in patients with atrial fibrillation. *Lancet* 2000 355: 956-962.
- 2 R Thomson et al. Guidelines on anticoagulant treatment for atrial fibrillation in Great Britain: variation in content and implications for treatment. *BMJ* 1998 316: 509-513.
- 3 IA Giannakakis et al. Citation of randomised evidence in support of guidelines of therapeutic and preventive interventions. *Journal of Clinical Epidemiology* 2002 55: 545-555.

SURGERY FOR ARTHRITIC KNEES

When medical therapy has failed to relieve pain of knee osteoarthritis, an additional treatment often used is lavage or debridement using arthroscopy. The evidence on which this has been based was a number of case series studies, showing that about half of patients so treated did reasonably well in the long term. A number of randomised trials using a variety of techniques and comparators have tested efficacy of the techniques. Most, but not all, claimed efficacy for lavage and/or debridement.

There are problems with these studies. All were small, some were of short duration, and some used surrogate end points like muscle strength. None used a true "placebo" control of sham operation. The accumulation of evidence that these procedures were beneficial is neither large, nor compelling.

So when a new randomised trial, superbly conducted and relatively large pops up and tells us these procedures do not work [1], we need to put our thinking caps on. We need to balance what we think we know with what someone is telling us we should know.

Randomised trial

Patients recruited were under 75 years of age with American College of Rheumatology criteria for osteoarthritis of the knee, with pain of ≥ 40 mm on a 100 mm scale (moderate), and with no arthroscopy for at least two years. They were randomised into three groups:

- ◆ Placebo: in which a standard arthroscopic debridement procedure was simulated, including skin incisions.
- ◆ Lavage: in which after diagnostic arthroscopy the joint was flushed with at least 10 litres of fluid.
- ◆ Debridement: in which after diagnostic arthroscopy the joint was flushed with at least 10 litres of fluid and cartilage shaved, loose debris removed, all torn or degenerated meniscus trimmed and the remaining meniscus smoothed to a firm and stable rim.

A single experienced surgeon who had been physician to the 1996 US Olympic basketball team carried out all the operations. Postoperative outcome assessments were made by other personnel, who, like the patients, were unaware of the procedure. Information on pain and other outcomes was collected for up to two years after the operation.

Results

Patients were predominantly men (over 90%), with a mean age of about 52 years, mostly taking non-prescription analgesics, but whose osteoarthritis was moderate or severe in about 70%. There were 60 per group at the start, with good follow up to two years.

At baseline the mean knee pain score was 65/100 mm in all groups, and was about 50/100 mm at one and two years in all three treatment groups. No outcome showed any difference between placebo and the two active treatments.

This was a trial with exemplary methods, showing conclusively that neither arthroscopy with lavage nor lavage plus debridement had any effect on pain or function. Yet among previous trials there was at least one [2] with very good methods, being properly randomised and with a blinded observer, and showing benefit for both intra-articular corticosteroid and lavage. It was small, though, with as few as 20 patients per group, and follow up was for only 24 weeks, by which time differences between groups were much less marked.

What were the differences between them? The American study [1] was in men, predominantly in their 50s. The French study [2] was mostly in women, and predominantly in their 60s. In both studies the initial pain intensities were about the same, and variability was great. Typically a mean pain score of 50 mm had a standard deviation of 20-30 mm. That means that the 95% confidence interval was 0-100 mm. **Bandolier** suspects, by analogy with other pain trials, that these data were not normally distributed, yet were reported as means. This would complicate things even more.

Size and bias

So what is the truth? Do we trust custom and practice, or high quality evidence? Is it all down to differences in patients selected and techniques, or is there something we can grasp? The answer may be size and avoidance of bias.

Trials of small size have been the norm until recent times, and a survey has shown that randomised trials with over 100 participants became the norm only very recently [3]. Small size is a real problem because random chance can dominate. And many of the older studies favouring arthroscopy had, in addition to small numbers, designs that might well have allowed significant bias in favour of it. Small studies with poor design, of short duration and with outcomes lacking relevance can fool us into thinking it works. Older trials that used non-treatment control groups showed that untreated patients improved just as much.

With this new trial what we can say is that on average there is no convincing evidence for any beneficial effect of arthroscopy with lavage and debridement. Proving conclusively the negative is a bit more difficult. But with an estimated 650,000 procedures costing \$3.5 billion each year in the USA alone, it certainly makes one think. **Bandolier** has ordered all the controlled trials it can find, and will abstract them on the Internet site in the coming months.

References:

- 1 JB Moseley et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *New England Journal of Medicine* 2002 347: 81-88.
- 2 P Ravaud et al. Effects of joint lavage and steroid injections in patients with osteoarthritis of the knee. *Arthritis & Rheumatism* 1999 42: 475-482.
- 3 S McDonald et al. Number and size of randomized trials reported in general health care journals from 1948 to 1997. *International Journal of Epidemiological Association* 2002 31: 125-127.

FAMILY/PARENTING INTERVENTIONS FOR DELINQUENCY

Repetitive and persistent antisocial behaviour in adolescents is probably familiar to all of us. Technically it is called conduct disorder, and affects between 1 and 3% (though it often seems much more common than that). Delinquency is a special category of this where children break the law. About 2% of children and adolescents enter the juvenile justice system each year, and a minority persist and account for a large amount of police and court time. A new systematic review tells us that family and parenting interventions can make a difference [1].

Systematic review

The review had a heroic search strategy looking at many databases and contacting experts. Included were randomised trials of family and/or parenting interventions for children or adolescents aged 10-17 with conduct disorder and/or delinquency. Conduct disorder was assessed by a standardised psychological checklist or psychiatric diagnosis. Delinquency was defined as referral to a justice system or similar for a serious crime (eg assault) and re-offending on at least two occasions. Sex and drug offences were excluded, as these are unlikely to be wholly related to conduct disorder. Only objective or validated outcomes were used.

Results

Eight trials were found with 749 children. Seven of the trials were blind and follow up was good. The follow up period varied between two months and four years. Interventions were usually quite intensive and the control was predominantly usual care. There was a predominance of boys in most studies, and two were 100% male.

Figure 1: Average institutional days in delinquent adolescents with family and parenting intervention or usual care

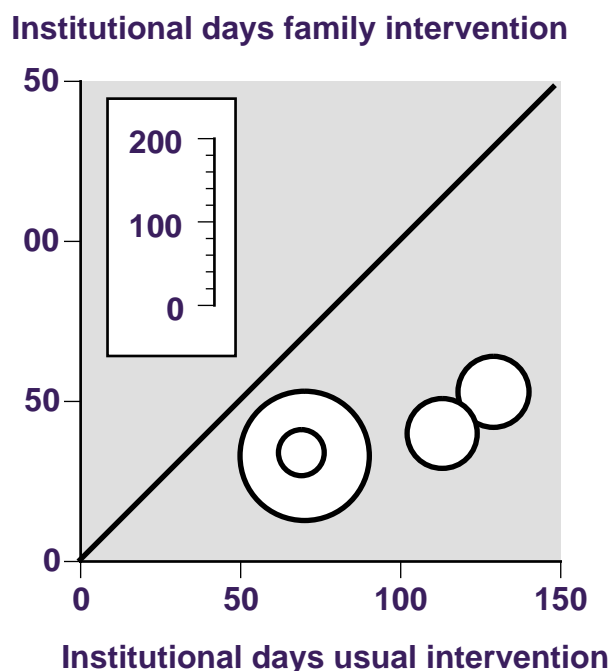
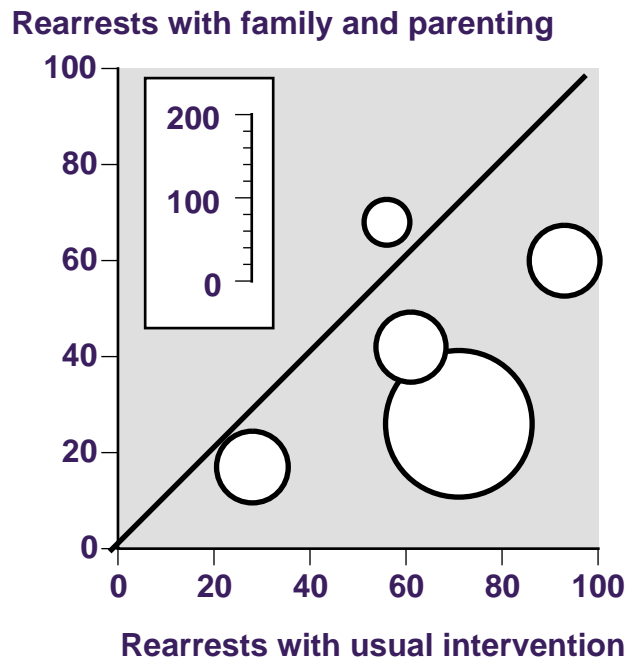


Figure 2: Re-arrest rates in delinquent adolescents with family and parenting intervention or usual care



Four studies (Figure 1) had information on days spent in an institution (prison, detention, community treatment centre). In each of these the average number of days was lower for the family and parenting intervention, with a weighted mean difference of 51 days and high statistical significance.

Five studies reported the number of delinquents who had been re-arrested during follow up. Family and parenting reduced the risk of re-arrest in four of the five studies (Figure 2). With usual care the re-arrest rate was 65% and with family and/or parenting interventions it was 37%. The relative risk was 0.66 (95% confidence interval 0.44 to 0.98) and the number needed to treat to prevent one adolescent delinquent being re-arrested was 3.7 (2.8 to 5.4). Higher quality studies gave a lower (better) NNT of 2.7.

The rate of subsequent arrests over one to three years was reported in five trials, with an average of nearly one fewer arrest in the treated group. Other outcomes were reported in few studies.

Comment

Most of us are lucky enough not to face these problems with adolescent children. The evidence here is that family and parenting interventions can help. There are hints that it may also benefit siblings. This is a bit outside drugs and surgery, but social issues affect health and wellbeing, and there are times when we need to know what to do to help.

References:

1 SR Woolfenden et al. Family and parenting interventions for conduct disorder and delinquency: a meta-analysis of randomised controlled trials. Archives of Diseases in Childhood 2002 86: 251-256.

SMOKING AND IMPOTENCE

One of the real problems with giving advice about smoking to young men is that we are telling them not to smoke now because of health problems they may experience decades in the future. Perhaps a more direct approach on cigarette packets would be more helpful. Not “smoking kills” but rather “smoking makes you impotent”. A new systematic review of smoking prevalence among impotent men points in that direction [1].

Systematic review

The review searched MEDLINE from 1980 to about 2000 for articles on impotence or erectile dysfunction and which described studies in the United States. The specific US choice was to be able to examine available smoking prevalence data for the general population. From each study the number of impotent men who were current smokers, the definition of impotence, the definition of smoking, ages, and location were taken.

For each study, a tailored comparison group based on age distribution, time and location was derived from smoking prevalence data from an ongoing surveillance system.

Results

There were 19 studies with 3,819 impotent men, ranging in size from 10 to 800 men. Studies usually had a wide age range, with mean age from 35 to 61 years.

In 16 of the 19 studies, smoking rates in impotent men were higher than those in the general male population (Figure 1). Overall the age and location matched smoking prevalence in men was 28%. In impotent men it was 40%, a significant difference of 12% (95% confidence interval 11 to 14%).

Figure 1: L'Abbé plot of smoking rates in impotent men and general male population

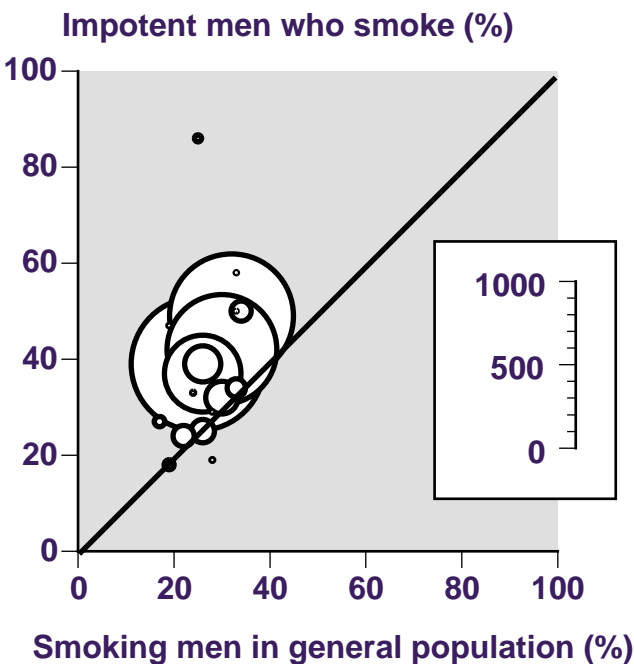
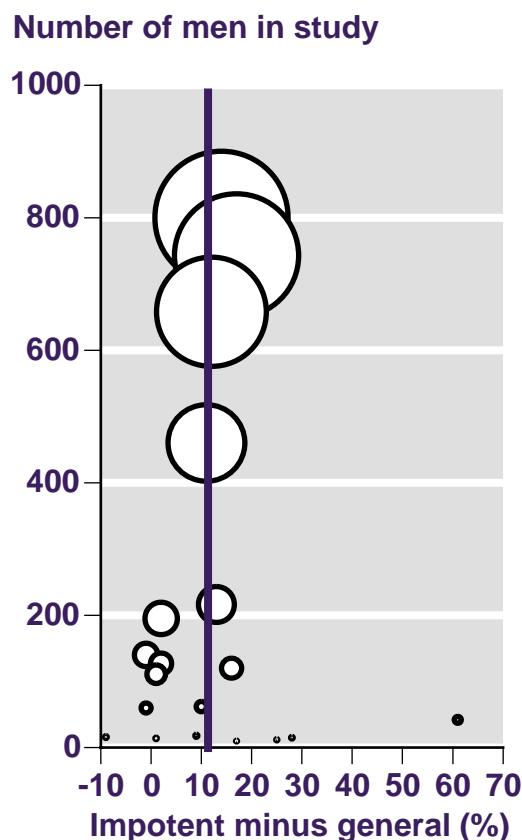


Figure 2: Difference in smoking prevalence in impotent men and general male population by study size



The difference between smoking rates in impotent men and the general male population of 12% was accurately measured by the larger, but not the smaller studies (Figure 2). With over 400 men the difference was consistent. With fewer than 100 men the difference ranged from 60% to -9%.

Comment

This paper alone does not constitute a definitive link between smoking and impotence in men. But it does go a long way to making the claim, and in the end it may be more influential in terms of making young men think twice about starting smoking, and make older men think about giving it up. Women may also appreciate the information.

There are at least two systematic reviews of smoking and erectile dysfunction published recently [2, 3]. The bottom line from these is that smoking doubles the risk of erectile dysfunction in men. Erectile problems are common in men, affecting 10-25%. All the evidence may not yet be in, but men might like to work on the basis that smoking will exacerbate a problem that age and chronic disease will bring to their door at some time.

References:

- 1 TO Tengs, ND Osgood. The link between smoking and impotence: two decades of evidence. *Preventative Medicine* 2001 32: 447-452.
- 2 KT McVary et al. Smoking and erectile dysfunction: evidence based analysis. *Journal of Urology* 2001 166:1624-32.
- 3 G Dorey. Is smoking a cause of erectile dysfunction? A literature review. *British Journal of Nursing*. 2001 10:455-65.

DEATH AND PROFIT

Most healthcare payment and delivery is through organisations with no profit motive. In some parts of the world, some hospitals are run on a 'for-profit' basis. A new systematic review suggests that these hospitals have higher mortality rates than not-for-profit institutions [1].

Review

This review is part of a larger study exploring the effects of the profit motive in healthcare. A complex search strategy examined multiple databases for studies examining hospital mortality in hospitals run for profit, and not for profit. Both observational and randomised studies were looked for. Mortality, types of patients, and adjustment for potential confounding criteria were among information extracted from the studies.

Results

Fifteen observational studies were included. All were in the United States, and in most patient care was publicly funded through Medicare. Most included medical, surgical and general acute care, with one specifically examining maternity services. The studies, between 1982 and 1995, looked at 26,000 hospitals, and with data on 38 million patients.

Fourteen studies evaluated adult populations. When adjusted for potential confounders, six had a significantly higher risk of death in for-profit hospitals, and one a significantly lower risk of death in for-profit hospitals. The pooled relative risk showed that private for-profit hospitals were associated with a higher risk of death. The relative risk was 1.020 (95% confidence interval 1.003 to 1.038). One study on perinatal mortality with 1.6 million patients in 243 hospitals also demonstrated an increased risk of death in private for-profit hospitals.

Comment

The first thing to take on board is that this review is not about ownership. Hospitals could be private or public. This is about whether the private hospitals were administered so that they had to make a profit. Those that did, compared with those that did not, had a higher death rate. Is the result credible? David Naylor [2] makes a good argument that it is. The excessive mortality is found in most of the studies, and there may be good reasons to treat conservatively the one study that went the other way. The reasons for the excess mortality are obvious. For-profit hospitals have to make a profit, and probably pay more to senior executives. The result is less resource at the front line.

References:

- 1 PJ Devereaux et al. A systematic review and meta-analysis of studies comparing mortality rates of private for-profit and private not-for-profit hospitals. *Canadian Medical Association Journal* 2002 166: 1399-1406.
- 2 CD Naylor. Your money and/or your life? *Canadian Medical Association Journal* 2002 166: 1416-1417.

DEPRESSION IN ADVANCED DISEASE

Depression in advanced disease is common, but might be easily overlooked. With an emphasis on improving the quality of life, it should not be. What we need is good evidence about how to spot it, and good evidence about how to treat it. Two systematic reviews from London and Stanford [1,2] make a start.

Depression prevalence

This systematic review [1] set out to find all studies examining the prevalence of depression in advanced disease. The problems were two: defining depression and defining advanced disease. For the first, the aim was to examine different methods of detecting depression, from clinical recognition to using different depression scales. For the second, reviewers had a target population of patients in palliative care settings, excluding studies focusing on non-cancer conditions like heart failure or AIDS. At least 50% of patients had to have cancer, and oncology studies were included if expected or real survival was less than six months.

A very extensive search strategy was used, examining many electronic databases, the St. Christopher's Hospice database, the Cochrane Library, and with handsearching selected journals.

Results

The review performs no meta-analysis to give an overall figure for depression, because there was real clinical heterogeneity between studies. The six main diagnostic criteria were clinical recognition, single-item scales, the Hospital Anxiety and Depression Scale (HADS), other questionnaires, use of diagnostic criteria and case-finding studies.

Figure 1 shows an abacus plot of the prevalence rates found for the first five of these methods in just under 6,700 patients in which the overall rate of depression was about 30%. For HADS and others reporting degrees of depression, only severe depression was included for Figure 1. Many studies were very small, and this also contributed to the wide variation, which was from under 10% to over 60%.

Comment

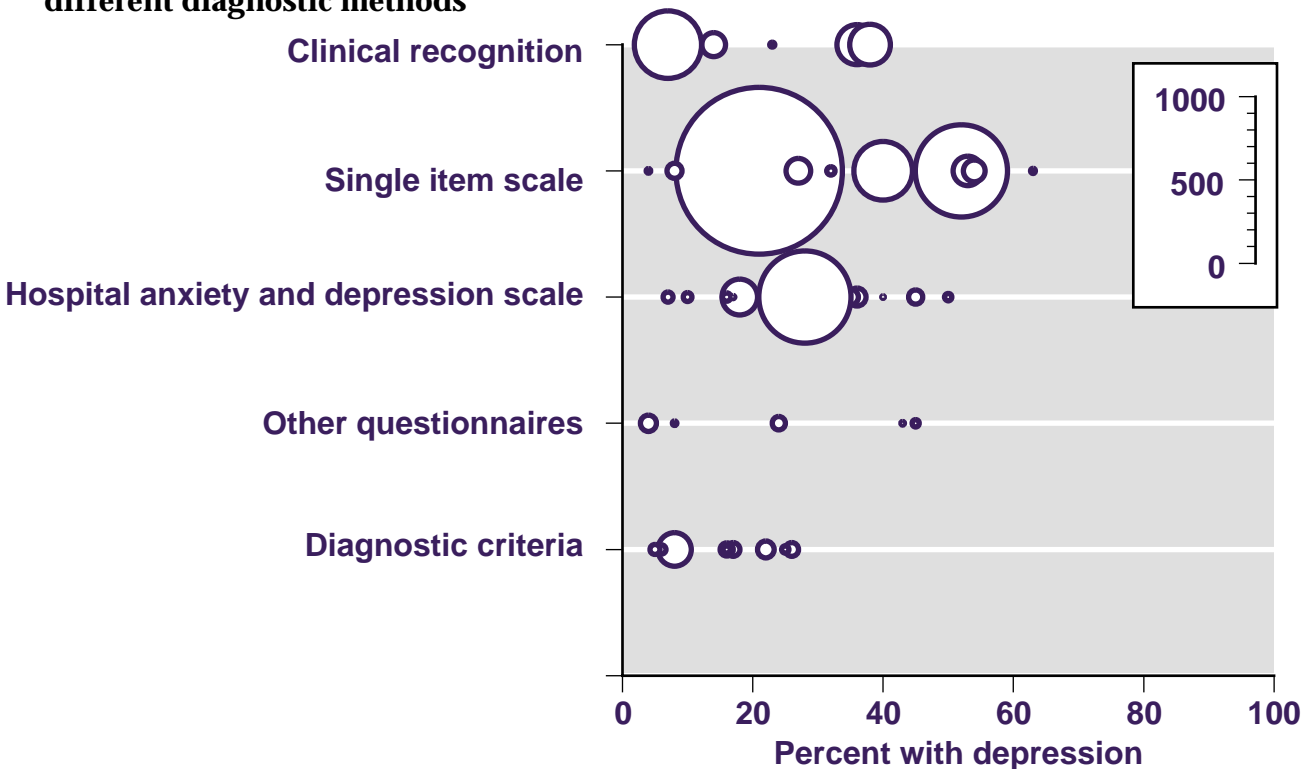
It is not possible to say that any particular method finds more depression than another, or which is correct. What is clear is that it is relatively common. A sister review on treatments [2] found no reports for psychosocial interventions, and only three randomised trials with antidepressants with a total of 163 patients. Two showed antidepressants to be better than placebo, and one showed fluoxetine to be as effective, but with lower discontinuations, than desipramine. The authors also comment on literature showing that antidepressants may be underused in a palliative care setting.

Palliative care has few randomised trials and even fewer systematic reviews. These constitute a significant addition to the knowledge, as opposed to the experience, base. Experience in palliative care is good, and has enabled great advances to be made. Knowledge will cement and improve that.

References:

- 1 M Hotopf et al. Depression in advanced disease: a systematic review part 1. Prevalence and case finding. *Palliative Medicine* 2002 16: 81-97.
- 2 KL Ly et al. Depression in advanced disease: a systematic review part 2. Treatment. *Palliative Medicine* 2002 16: 279-284.

Figure 1: Prevalence of depression in patients with advanced disease according to different diagnostic methods



INFLUENZA VACCINE IN OVER-65s

Influenza vaccination time is coming round again. Many of the arguments have been won, and many older people living at home now have their annual jab. Not all do, and not everyone is convinced. In *Bandolier* 11 we examined a randomised trial of influenza vaccination in Holland, and in *Bandolier* 73 a review on people younger than 65 years. Both were persuasive. A new systematic review extends the evidence base [1].

Systematic review

In that wonderful city of Melbourne, the authors used an extensive search strategy for studies in any language where inactivated influenza vaccine was used, the study period was determined by influenza surveillance, and the population was community-based elders of 65 years or more. Excluded were studies on institutionalised persons, where comparability of study groups was not reported, studies with a cross-sectional design and those where influenza strains used did not match circulating strains, or where this was not reported.

The outcome used was vaccine effectiveness. Vaccine effectiveness (VE) is given by the equation:

$$VE = (1 - \text{relative risk}) \times 100$$

Thus a relative risk of 0.5 in an outcome gives a vaccine effectiveness of 50%, or a 50% reduction in the rate of that outcome over control.

Results

There were 15 studies. One was a randomised trial (*Bandolier* 11), two others were labelled as clinical trials, and the remainder were case-control or cohort studies. Results of vaccine effectiveness for different outcomes are shown in Figure 1. The biggest effect was on mortality.

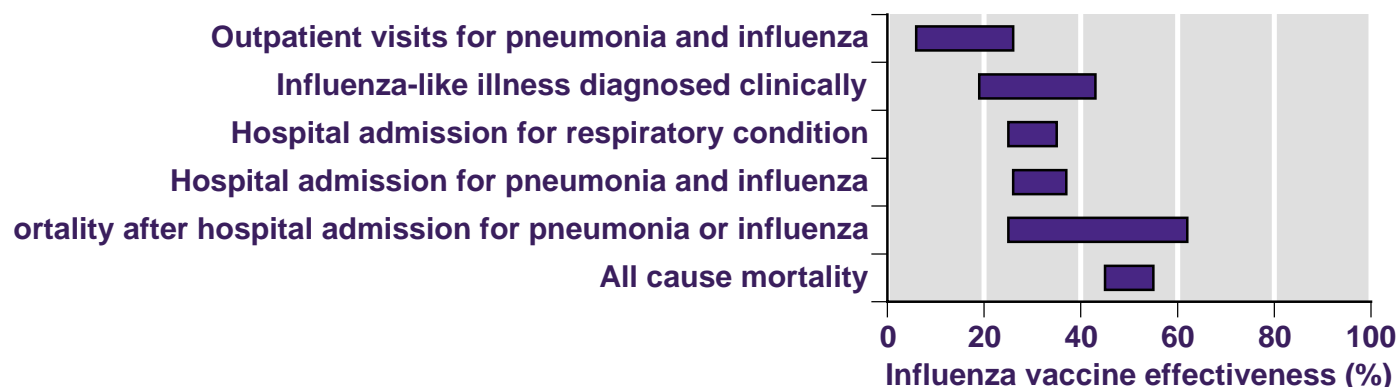
Comment

This confirms the effectiveness of influenza vaccines for a variety of outcomes. The results confirm other meta-analyses from healthy adults or institutionalised elderly. So why bother doing the study? The answers to this are several, but the main importance to us is likely to be a reaffirmation that something done every year is worth doing, and worth doing well.

References:

- 1 T Vu et al. A meta-analysis of effectiveness of influenza vaccine in persons aged 65 years and over living in the community. *Vaccine* 2002 20: 1831-1836.

Figure 1: Vaccine effectiveness of influenza vaccine in the over 65s measured by different outcomes



Subscribing to Bandolier

Readers who would like their own copy have several ways of subscribing. You can:

Download a subscription form from the Internet (www.jr2.ox.ac.uk/bandolier/backnos.html) and fax it to Maura Moore on +44 1865 226978

Contact Maura for a subscription form by telephone (+44 1865 226132), fax (+44 1865 226978), or by email to maura.moore@pru.ox.ac.uk

Subscription rates are £36 in the UK and £72 overseas.

EDITORS

Andrew Moore Henry McQuay

Pain Relief Unit
The Churchill, Oxford OX3 7LJ

Editorial office: 01865 226132

Editorial fax: 01865 226978

Email: andrew.moore@pru.ox.ac.uk

Internet: www.ebandolier.com

ISSN 1353-9906